

# Size-frequency distribution and the large deviation function for frequency in simple models of earthquakes: a scaling approach

T. Mitsudo\* and N. Kato

*Earthquake Research Institute, the University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo, Japan.*

(Dated: September 6, 2012)

The statistical properties of large, disastrous earthquakes have significance in the study of deviations from the regular behavior of earthquakes. In this study, to help understand the irregular behavior of earthquake occurrences, the large deviation function for the frequency of earthquakes is calculated for the first time. We study the temporal sequence of the largest earthquakes in simple models of earthquakes, one-dimensional forest-fire models in which the fluctuation in loading and fracture process are taken into consideration. We introduce four different models of fixed trigger sites to represent the starting points of rupture propagation. The size-frequency distributions and scaled large deviation functions for the frequency of the largest earthquakes in the system are calculated and the phase tables are presented. The calculated large deviation functions are compared with those of the homogeneous Poisson process and of the one-site forest-fire model. We find that the large deviation function largely depends on the model parameters and the fixed trigger sites, and in most cases the large deviation function deviates from that of the homogeneous Poisson process, suggesting that the largest earthquakes in the model do not occur randomly. The relation between the size-frequency distribution and the large deviation function for the frequency is discussed.

## I. INTRODUCTION

Statistical indices for characterizing earthquakes are important for understanding the mechanism of earthquakes. The Gutenberg-Richter (GR) law [1] for the magnitude of an earthquake is well established in seismology, and it is well known that the  $b$ -value depends on both time and space. However, in the tail-part of the distribution that corresponds to large magnitudes, the functional form is still a subject of concern [2] because we have insufficient data from large earthquakes. Moreover, being able to estimate the frequency or probability of rare events such as disastrous earthquakes is a big concern from the point of view of hazard assessment.

To further study the frequency of large earthquakes, we adopted a large deviation function (LDF) [3, 4]. The LDF is related to the probability of rare events, which constitute the tail part of the probability distribution. The LDF is universal in the sense that it is an asymptotic form whenever the number of elements is large. Recently, much interest has been paid to the LDF for current in nonequilibrium statistical physics, where the LDF is expected to serve as a “thermodynamic” function [5, 6]. The LDF is also used in the analysis of activity in glassy systems, at the transition between active state and inactive state [7]. Besides the LDFs for current and activity, the LDF for frequency is also of recent concern in counting processes such as photon-counting processes [8–10]. Budini [9] studied the “thermodynamic” framework of a point process by using the LDF for temporal frequency. The earthquake sequence can be regarded as a point process such as the well-known epidemic-type aftershock sequence (ETAS) model [11]. The “thermodynamic” approach can be applied to such earthquake point processes. Studies that analytically calculated the LDF in “thermodynamics” and statistical physics, e.g., [5], preceded numerical calculations, even though analytical calculation of the LDFs in complex systems is difficult. Recently, Monte Carlo methods have been introduced to obtain the LDF of a system described by master equations in discrete time [12] and in continuous time [13, 14]. These methods enabled us to calculate the LDFs in the complex systems of earthquakes.

Because the recurrence time of large earthquakes is longer than several tens of years, we do not have enough data on large earthquakes, and observational studies of the frequency of large earthquakes are limited. In order to overcome this, we can generate enough data by simulating earthquakes. In this paper, we study the LDF for the frequency of simulated earthquakes in a one-dimensional forest-fire model, which can be understood as a minimalist model for earthquakes. Originally, the forest-fire model was introduced to simulate forest-fires [15]. Drossel and Schwabl [16] represented forest-fires in the model with four processes: planting of trees, ignition, propagation of fire, and extinguishing of fires. To separate the timescale of the planting process and those of the latter three, a ‘lightning’ model was introduced [17, 18]. The lightning model reduces the last three processes to a single process, the vanishing of a cluster of trees.

---

\*Electronic address: mitsudo@eri.u-tokyo.ac.jp

Besides recent applications to real forest-fires [19], forest-fire models have been applied to simulate earthquakes [20]. When used as an earthquake model, loading on a fault corresponds to planting a tree and triggering of an earthquake corresponds to igniting a fire. The idea of representing an earthquake with a randomly expanded cluster was introduced by Otsuka [21]. Newman and Turcotte [22] studied the cycles of large earthquakes, which were defined as the percolated clusters in a 2D forest-fire model. To establish a minimalist model of earthquakes and to estimate the predictability of earthquakes, Vázquez-Prada et al. [23] arrived at a model similar to a 1D lightning forest-fire model. Recently, heterogeneous configurations of ignition (trigger) sites were introduced by Tejedor et al. [24] in 2D forest-fire models to represent the variation of faults. They found that the size-frequency distributions of the simulated earthquakes could be expressed by power laws. In their models, there are some asperity regions where large earthquakes tend to occur. According to the behavior of the size-frequency distribution, they obtained phase maps, which were found to be dependent on the configuration of the trigger sites.

In this study, we adopted four 1D forest-fire models with different numbers of trigger sites as simple models of earthquakes. We obtained the size-frequency distribution and numerically calculated the LDFs for the frequency of the system-size earthquakes. A system-size earthquake is an earthquake whose size is characterized by the system size of a model, and it is the largest earthquake in the system. We classified the behaviors of the size-frequency distribution and the LDFs for frequency of the system-size earthquakes into ‘phases’. The LDFs for frequency of the system-size earthquakes were scaled by the mean frequency of the system-size earthquake.

First, we introduce the four models used in this study and write their master equations. We give a brief introduction to the LDF used in this study and to the method of calculating the LDF. Next, we show the results of the size-frequency distributions of the four models and present a phase table. Similarly, the LDFs for the frequency of the system-size earthquakes scaled by the mean frequency are calculated to obtain a phase table. We discuss the relationship between the phase of the size-frequency distribution and that of the LDF, and we present our conclusions.

## II. MODEL AND LARGE DEVIATION FUNCTION

### A. One-dimensional forest-fire models

We studied the forest-fire models in 1D on a lattice of length  $L$ . To take into account the heterogeneous nature of faults, we introduced the four models as shown in Fig. 1, M1, M2, M3, and MA. M1 has trigger sites only at the left edge, M2 can have triggers at both the edges, M3 can have triggers at both the edges and at the site  $m$  ( $2 \leq m \leq L-1$ ), and MA can have triggers at all sites. MA is a common homogeneous forest-fire model [25], while M1, M2, and M3 may represent heterogeneous faults. In M1, M2, and M3, earthquakes of various sizes can nucleate at the trigger sites, and the other sites are broken only by large earthquakes which nucleated at the trigger sites.

To simulate a cellular automaton model such as a forest-fire model, an update rule is required [26]. Among the cellular-automaton-like models, the results of a random update rule are consistent with steady state solutions [27, 28] of the master equation in continuous time for an asymmetric simple exclusion process [29]. For the lightning model, a formulation of the master equation in continuous time is available [30], and this enables us to calculate numerically the LDF. Thus we adopted the random update rule in the models M1, M2, M3, and MA.

The simulation procedure is as follows: First a site is chosen at random. If the site is empty, the site is loaded at the rate  $p$  per unit time. If the site is loaded and the chosen site is a trigger site, an earthquake is triggered at the rate  $f$  per unit time. When an earthquake is triggered, triggering also occurs at the neighboring sites until an unloaded site is encountered. The probabilities of the loading and the triggering processes describe the fluctuation of the dynamics in the loading and rupturing processes.

The state of the  $j$ th site ( $j = 1, \dots, L$ ) is described by the occupation number  $\tau_j$ , which takes 1 when the site is loaded and 0 when unloaded. Loading stress on the  $j$ th site is expressed by the transition from  $\tau_j' = 0$  to  $\tau_j = 1$ , and stress release (an earthquake) by the transition from  $\tau_j' = 1$  to  $\tau_j = 0$ , where  $'$  represents the configuration before the transition. When an earthquake is triggered at one of the loaded sites from  $j+1$  to  $j+s$ , and the sites  $j$  and  $j+s+1$  are empty, an earthquake of size  $s$  occurs, and it is expressed by the transition from  $\{\tau_j', \dots, \tau_{j+s+1}'\} = \{0, 1, \dots, 1, 0\}$  to  $\{\tau_j, \dots, \tau_{j+s+1}\} = \{0, \dots, 0\}$ .

For simplicity, we write the configuration of the system  $C = \{\tau_1, \dots, \tau_L\}$ . We introduce a probability  $P(C; t)$  that the system is in the configuration  $C$  at time  $t$ . The master equation is written as

$$\frac{dP(C; t)}{dt} = \sum_{C' \neq C} [W(CC')P(C'; t) - W(C'C)P(C; t)], \quad (1)$$

where  $W(CC')$  is the transition rate from  $C'$  to  $C$ . The summation  $\sum_{C' \neq C}$  represents the sum over all configurations

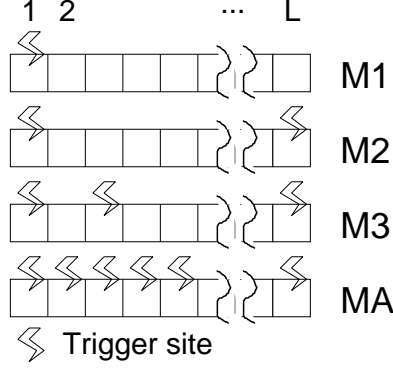


FIG. 1: Description of models M1, M2, M3, and MA with lattices of length  $L$ . Trigger sites are represented by lightning marks.

of  $C'$  except for  $C$ .  $W(CC')$  is written as

$$W(CC') = p \sum_{j=1}^L [\cdots \delta_{\tau_{j-1}\tau'_{j-1}} (1 - \tau'_j) \tau_j \delta_{\tau_{j+1}\tau'_{j+1}} \cdots] + W_f(CC'), \quad (2)$$

where  $W_f(CC')$  is the trigger term that depends on the model, and  $\delta_{xx'}$  is the Kronecker delta. The parts with a suffix less than 1 or greater than  $L$  are replaced by unity (e.g.,  $\delta_{\tau_0\tau'_0}$  is replaced by 1). The exact forms of  $W_f(CC')$  are written as

$$W_f(CC') = f \sum_{j=1}^L \tau'_1 (1 - \tau_1) \cdots \tau'_j (1 - \tau_j) (1 - \tau'_{j+1}) (1 - \tau_{j+1}) \delta_{\tau_{j+2}\tau'_{j+2}} \cdots \quad (3)$$

for M1,

$$\begin{aligned}
 W_f(CC') = W_2 = & f \sum_{j=1}^L \tau'_1 (1 - \tau_1) \cdots \tau'_j (1 - \tau_j) (1 - \tau'_{j+1}) (1 - \tau_{j+1}) \delta_{\tau_{j+2}\tau'_{j+2}} \cdots \\
 & + f \sum_{j=1}^L \tau'_{L-j} (1 - \tau_{L-j}) \cdots \tau'_L (1 - \tau_L) (1 - \tau'_{L-j-1}) (1 - \tau_{L-j-1}) \delta_{\tau_{L-j-2}\tau'_{L-j-2}} \cdots
 \end{aligned} \quad (4)$$

for M2,

$$\begin{aligned}
 W_f(CC') = & W_2 + f \sum_{j=1}^m \sum_{k=m}^L \cdots \delta_{\tau_{j-2}, \tau'_{j-2}} [(1 - \tau'_{j-1}) (1 - \tau_{j-1}) \tau'_j (1 - \tau_j) \cdots \\
 & \tau'_k (1 - \tau_k) (1 - \tau'_{k+1}) (1 - \tau_{k+1})] \delta_{\tau_{k+2}\tau'_{k+2}} \cdots
 \end{aligned} \quad (5)$$

for M3, and

$$\begin{aligned}
 W_f(CC') = & W_2 + f \sum_{m=2}^{L-1} \sum_{j=1}^m \sum_{k=m}^L \cdots \delta_{\tau_{j-2}, \tau'_{j-2}} [(1 - \tau'_{j-1}) (1 - \tau_{j-1}) \tau'_j (1 - \tau_j) \cdots \\
 & \tau'_k (1 - \tau_k) (1 - \tau'_{k+1}) (1 - \tau_{k+1})] \delta_{\tau_{k+2}\tau'_{k+2}} \cdots
 \end{aligned} \quad (6)$$

for MA.

## B. Large deviation function

The mean frequency of earthquakes of size  $s$  per unit time  $x(s)$  is written as

$$x(s) = \frac{N(s)}{t}. \quad (7)$$

Here,  $N(s)$  is the number of earthquakes of size  $s$  for elapsed time  $t$ . We may write  $N(s)$  as

$$N(s) = \sum_{i=1}^{N_T} X(s; C_{i+1}, C_i), \quad (8)$$

where  $X(s; C_{i+1}, C_i)$  is 1 when an earthquake of size  $s$  occurs and is 0 otherwise, when the configuration changes from  $C_i$  to  $C_{i+1}$ , and  $N_T$  is the total number of configuration changes. The probability that the mean value  $x(s)$  is equal to the frequency  $x$  is written as  $P(x)$ .  $P(x)$  is asymptotically written as

$$P(x) \sim e^{-t\phi(x)} \quad (9)$$

for large  $t$ , where the function  $\phi(x)$  is called the large deviation function (LDF). The LDF satisfies  $\phi(x_m(s)) = 0$ , which indicates the statistical mean  $x(s)$  converges to the true mean  $x_m(s)$  when  $t \rightarrow \infty$ , and  $\phi(x) \geq 0$ . For earthquakes,  $X(s; C_2, C_1), \dots, X(s; C_{N_T}, C_{N_T-1})$  are not independent and identically distributed (i.i.d.) random variables. For i.i.d. random variables we have the central limit theorem. In contrast, when we study earthquakes, the LDF becomes significant because of the non-i.i.d. nature.

A large deviation function  $\phi(x)$  has a corresponding ‘generating function’  $\mu(\lambda)$ , where  $\lambda$  is the conjugate variable of  $x$ . The generating function  $\mu(\lambda)$  is defined as

$$e^{\mu(\lambda)t} = \langle e^{\lambda x t} \rangle \sim \int e^{(\lambda x - \phi(x))t} dx. \quad (10)$$

$\phi(x)$  and  $\mu(\lambda)$  are related by the Legendre transform as

$$\phi(x) = \max_{\lambda} [x\lambda - \mu(\lambda)]. \quad (11)$$

For example, the large deviation function for the frequency of events in a point process that obeys a homogeneous Poisson process is

$$\phi_P(x) = x \log\left(\frac{x}{\alpha}\right) - x + \alpha, \quad (12)$$

and the corresponding generating function is

$$\mu_P(\lambda) = \alpha(e^\lambda - 1), \quad (13)$$

where  $\alpha$  is the rate of event occurrence and the suffix  $P$  represents the Poisson process.

To estimate the generating function  $\mu(\lambda)$ , we calculate the modified master equation written as

$$\frac{d}{dt} P^\lambda(C; t) = \sum_{C' \neq C} [W^\lambda(CC') P^\lambda(C'; t) - W(C'C) P^\lambda(C; t)], \quad (14)$$

where  $P^\lambda(C; t)$  satisfies the differential equation (14) under the initial condition  $P^\lambda(C_0; t_0) = P(C_0; t_0)$  with an initial configuration  $C_0$ , and  $W^\lambda(CC') = W(CC')e^{\lambda X(s; C, C')}$ . This modified master equation is also expressed by a matrix, and the largest eigenvalue of the matrix is asymptotically equal to  $\mu(\lambda)$  when  $t$  is large. We show here the calculation of the one-site forest-fire model to illustrate the derivation of  $\mu(\lambda)$ . The transition matrix of the one-site forest-fire model  $\mathbf{H}$  is written as

$$\mathbf{H} = \begin{pmatrix} -W(\{1\}\{0\}) & W(\{0\}\{1\}) \\ W(\{1\}\{0\}) & -W(\{0\}\{1\}) \end{pmatrix} = \begin{pmatrix} -p & f \\ p & -f \end{pmatrix}, \quad (15)$$

and the modified transition matrix  $\mathbf{H}^\lambda$  as

$$\mathbf{H}^\lambda = \begin{pmatrix} -W(\{1\}\{0\}) & W(\{0\}\{1\})e^{\lambda X(1; \{0\}, \{1\})} \\ W(\{1\}\{0\}) & -W(\{0\}\{1\}) \end{pmatrix} = \begin{pmatrix} -p & fe^\lambda \\ p & -f \end{pmatrix}. \quad (16)$$

By calculating the largest eigenvalue of this modified matrix, the generating function of the one-site system  $\mu_1(\lambda)$  is given by

$$\mu_1(\lambda) = \frac{1}{2} \left[ -p - f + \sqrt{(p - f)^2 + 4pfe^\lambda} \right]. \quad (17)$$

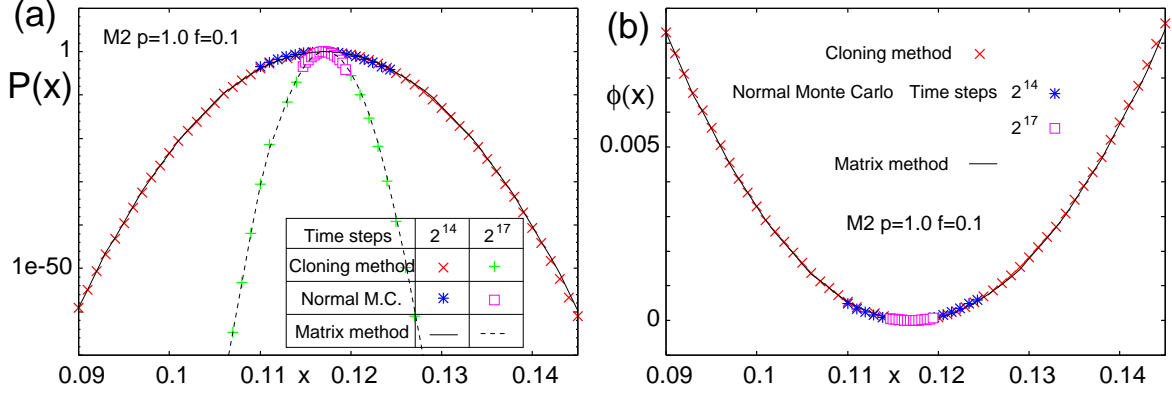


FIG. 2: (a) Comparison of  $P(x)$  obtained by the normal Monte Carlo (M.C.) method, the cloning Monte Carlo method, and the matrix method, with different time steps. (b) Comparison of  $\phi(x)$  obtained by the same methods as in (a).

The LDF is calculated from the generating function using the Legendre transform relation (11). By estimating  $\lambda^*$  that satisfies  $\frac{\partial}{\partial \lambda}(x\lambda - \mu(\lambda))|_{\lambda=\lambda^*} = 0$ , the LDF for the frequency of earthquakes of the one-site system  $\phi_1(x)$  is obtained as

$$\phi_1(x) = x \log \left[ \frac{x}{pf} (2x + \sqrt{4x^2 + (p-f)^2}) \right] + \frac{p+f}{2} - \frac{1}{2} \sqrt{(p-f)^2 + 8x^2 + 4x\sqrt{4x^2 + (p-f)^2}}. \quad (18)$$

The mean frequency of earthquakes in the one-site system is  $x_{1m}(1) = \frac{pf}{p+f}$ . This is the matrix method for obtaining the LDF.

In the present study, the largest eigenvalue of the modified transition matrix is calculated numerically for  $L < 14$  to obtain the generating functions. A cloning Monte Carlo method can also be used to calculate the generating function [13] for larger system sizes, where the matrix method cannot be used because the requirements of memory size and computation time would be unrealistic. However, the generating function is calculated more accurately by the matrix method than by the cloning Monte Carlo method. Details of the cloning Monte Carlo method used in this study are given in several references [13, 14, 31], so we will just give an outline of the method. For a given initial state, we prepare  $N_C$  copies (clones), each with its own clock. In each transition step, the earliest clone is chosen, and it evolves to a new configuration at a rate proportional to the modified transition rate. The amount of time that has elapsed from the previous transition and the number of the chosen clone would be copied or pruned ( $y$ ) are calculated after the transition. During the copying and pruning process in this procedure, the number of clones  $N_C$  is kept constant. The generating function is calculated iteratively from the values  $N_C$  and  $y$ .

The cloning Monte Carlo method and the matrix method can calculate the LDF in a wider range of  $x$  than can the normal Monte Carlo method, which is the method we used to simulate the forest-fire model. Values of  $P(x)$  and  $\phi(x)$  calculated by the three different methods are shown in Fig. 2 for the model M2 with  $p = 1.0$ ,  $f = 0.1$ , and  $L = 12$ . For the normal Monte Carlo,  $2^{14}$  and  $2^{17}$  time steps and  $2^{14}$  ensemble members were simulated, and  $P(x)$  was obtained by making a histogram of  $x$ .  $2^{11}$  clones were used in the cloning Monte Carlo.  $P(x)$  and  $\phi(x)$  of the normal Monte Carlo with  $2^{14}$  time steps exist for  $0.11 \leq x \leq 0.125$ , which is shown in Fig. 2, and the simulation result for the other values of  $x$  does not exist within the ensemble of the present calculations. This finite limiting range becomes narrower with increasing time steps for the normal Monte Carlo. The plotted range for the  $2^{17}$  time steps is narrower than that for the  $2^{14}$  time steps. Fig. 2 clearly shows that the results obtained from the three methods are in close agreement. This result indicates that the methods are giving sufficiently accurate values that are mutually consistent.

### III. SIMULATION RESULTS

#### A. Size-frequency distributions

In this study, the loading rate  $p$  is fixed 1.0, which corresponds to the rescaling of time by  $p$ . Fig. 3 shows the size-frequency distributions of simulated earthquakes for the models M1, M2, M3, and MA, with  $L = 128$ . The trigger

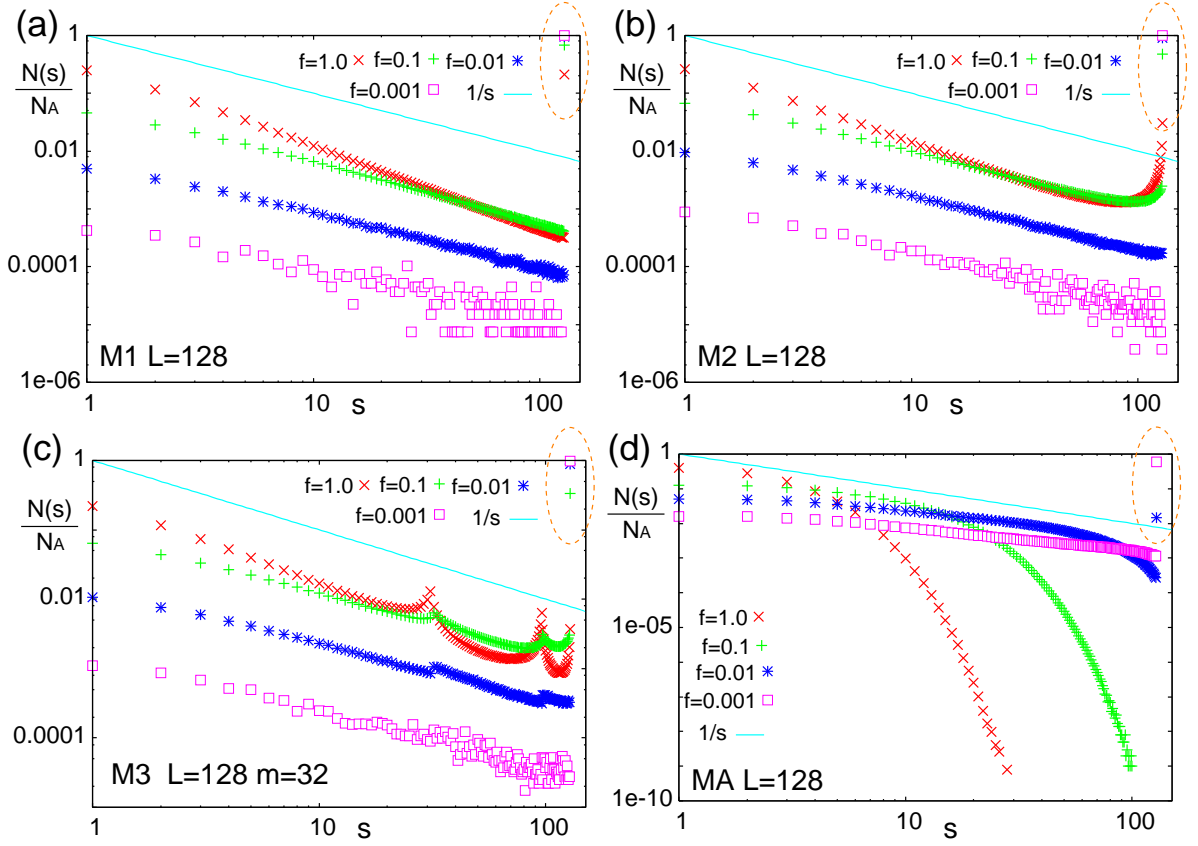


FIG. 3: Frequency-size distribution of earthquakes for the models M1, M2, M3, and MA with  $L = 128$ ,  $p = 1.0$ ,  $f = 1.0, 0.1, 0.01$ , and  $0.001$ . The solid line denotes  $1/s$ . The system-size earthquakes are highlighted by the orange dotted circles.  $N_A$  is the total number of earthquakes.

sites for M3 were located at sites 1, 32, and 128. The trigger rate  $f$  varies as 1.0, 0.1, 0.01, and 0.001. We took  $2^{27}$  Monte Carlo steps and recorded the number of earthquakes by size. The solid line in each graph of Fig. 3 denotes the function  $1/s$ , where  $s$  is the earthquake size. The frequencies of the system-size earthquakes are significantly large, with peaks at  $s = L$ , except for MA with  $f = 0.1$  and  $1.0$ .

Previous studies have found three types of behaviors ('phases') in the size-frequency distribution in the spring-block models and cellular automaton models of earthquakes [2, 24, 32, 33]. Following them, we classify the size-frequency distributions of the simulated earthquakes into three phases as follows: When the size-frequency distribution, excepting the system-size earthquakes, can be expressed by a power-law (the GR law), the behavior of the distribution is called 'critical'. When the frequencies of large earthquakes whose sizes are close to  $L$  are higher than would be expected by the power law fit to the smaller earthquakes, it is 'supercritical'. When the frequencies of large earthquakes whose sizes are close to  $L$  are lower than those expected by the power law, it is 'subcritical'. These behaviors in the size-frequency relations of earthquakes are observed in different areas and in different faults[2], though the phase description of the size-frequency distribution is not commonly used by seismologists.

The phases of the size-frequency distributions thus classified are summarized in Table I. For M1 (Fig. 3(a)), the size-frequency distributions show the critical phase, independent of  $f$ . The peak of the frequency at  $s = L$  is more significant for smaller  $f$ . This is because all the sites tend to be loaded before triggering for small  $f$ . The exponent of the decay of frequency with increasing  $s$  is approximately  $-1.0$  for  $f = 0.1$  and  $0.01$ . The decay rate is larger for larger  $f$ , because smaller earthquakes occur more easily for larger  $f$ .

For M2 (Fig. 3(b)), the supercritical phase is observed for  $f = 1.0$  and  $0.1$ , where the frequency of earthquakes increases with  $s$  for  $s \gtrsim 100$ , while the critical phase is observed for  $f = 0.01$  and  $0.001$ . Significant peaks at  $s = L$  are observed for all the values of  $f$ . For M3 (Fig. 3(c)), we observe peaks at  $s = 32, 96$ , and  $128$  for  $f = 1.0, 0.1$ , and  $0.01$ . These peaks correspond to the distances between the trigger sites. The peaks at  $s = 32$  and  $96$  tend to be unclear for smaller  $f$ , because more sites tend to be loaded before triggering and so the system-size earthquakes become prominent. The peaks of frequencies at  $s = 32$  and  $96$  in M3, and the supercritical behavior for M2 and M3,

	f=1.0	f=0.1	f=0.01	f=0.001
M1	<i>critical</i>	<i>critical</i>	<i>critical</i>	<i>critical</i>
M2	<i>super</i>	<i>super</i>	<i>critical</i>	<i>critical</i>
M3	<i>super</i>	<i>super</i>	<i>super</i>	<i>critical</i>
MA	<i>sub</i>	<i>sub</i>	<i>sub</i>	<i>sub</i>

TABLE I: Phases of the size-frequency distributions in Fig. 3, where *super*, *sub*, and *critical* denote the supercritical, subcritical, and critical phases, respectively.

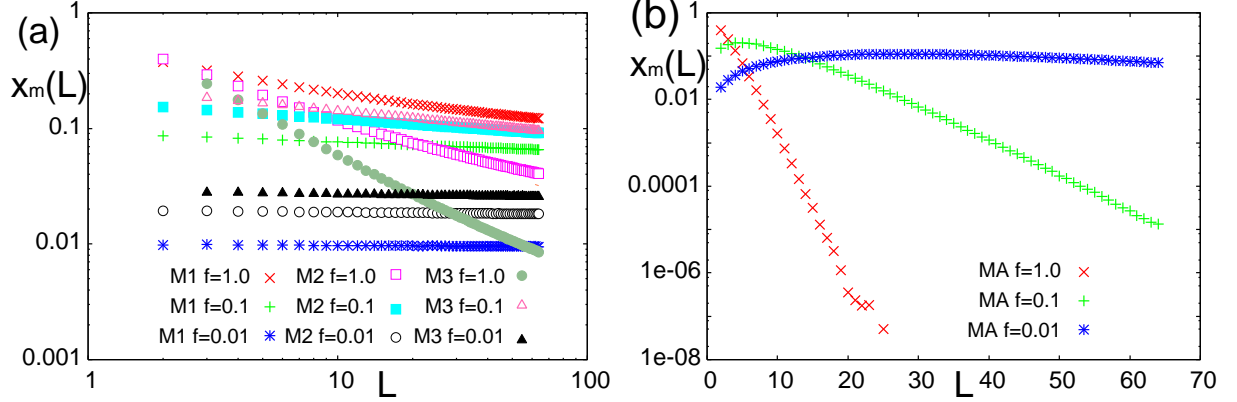


FIG. 4: Mean frequency of system-size earthquakes  $x_m(L)$  versus the system size  $L$ . (a) Models M1, M2, and M3, with  $f = 1.0, 0.1$ , and  $0.01$ . (b) Model MA with  $f = 1.0, 0.1$ , and  $0.01$ .

may be explained as follows: Small earthquakes occur frequently and release stress near the trigger sites, and clusters of loaded sites tend to be generated between the trigger sites. These clusters correspond to high frequencies of large earthquakes.

For MA (Fig. 3(d)), the subcritical phase is observed for all cases. The size-frequency distributions may be approximated by power laws for small  $s$ . The frequency decreases more rapidly with an increase in  $s$  than would be expected from the power law, and this rapid decrease starts at smaller values of  $s$  for larger values of  $f$ . The peak of the frequency of earthquakes of  $s = L$  was found for  $f = 0.01$  and  $0.001$ , although it was not found for  $f = 1.0$  and  $0.1$ .

### B. Scaled LDF of system-size earthquakes

Next, we calculated the LDFs for frequencies of the system-size earthquakes for the models M1, M2, M3, and MA, varying the system size  $L$  and the triggering rate  $f$ . Because the LDF takes the minimum value at the true mean frequency of the system-size earthquakes  $x_m(L)$ , we first evaluated  $x_m(L)$  for each case of simulation.  $x_m(L)$  depends on  $f$ ,  $L$ , and the choice of model (M1, M2, M3, or MA). Fig. 4 shows  $x_m(L)$  versus the system size for  $f = 1.0, 0.1$ , and  $0.01$ , and for models M1, M2, M3, and MA. For M3, the trigger sites are located at both the ends and at  $m = L/2 + 1$  for even  $L$  and  $m = (L + 1)/2$  for odd  $L$ . Note that the results for M1, M2, and M3 are plotted on logarithmic coordinates (Fig. 4(a)), while those for MA are on semilogarithmic coordinates (Fig. 4(b)). The time step was taken as  $2^{24}$  in each case. For M1, M2, and M3, with  $f = 1.0$  and  $0.1$ ,  $x_m(L)$  decreased with  $L$ . When  $f = 0.01$ ,  $x_m(L)$  seems to be independent of  $L$ , as shown in Fig. 4(a), although it approximately obeys power-law decay with an exponent of about  $-0.01$  for M1. For MA and  $f = 1.0$ ,  $x_m(L)$  exponentially decreases with increasing  $L$ . For  $f = 0.1$  and  $0.01$ ,  $x_m(L)$  takes the maximum values at  $L \sim 5$  and  $L \sim 27$ , respectively, and exponentially decreases with  $L$  for large  $L$ . The decrease of  $x_m(L)$  with an increase in  $L$  occurs because the preparation time for a system-size earthquake increases with  $L$ .

Because  $x_m(L)$  depends on  $L$ , appropriate scaling is required to compare the LDFs for different system sizes. To introduce scaling by  $x_m(L)$ , which is simply written as  $x_m$  hereinafter, we define a scaled variable  $z$  as

$$z = \frac{x - x_m}{x_m}. \quad (19)$$

In this scaling, the LDF of the Poisson process is given by

$$\frac{\phi_P(z)}{x_m} = (z + 1) \log(z + 1) - z. \quad (20)$$

The division by  $x_m$  on the left-hand side of (20) is unnecessary if the time  $t$  is scaled by  $x_m$ , and the LDF written by  $z$  omits  $x_m$  and can simply be written as  $\phi_P(z)$ . The introduction of  $z$  enables us to compare the calculated LDF to the LDF of the Poisson process (20) more clearly. We used two methods to calculate  $x_m$ , the Monte Carlo method, which was used for the results shown in Fig. 4, and fitting with the least-squares method. We prefer the least-squares method because we want the result to be self-consistent. To obtain  $x_m$  by the least-squares method, for each case, we numerically calculated  $\phi(x)$  and fit it with a polynomial, then used the  $x$  that minimizes this polynomial as  $x_m$  to scale the LDF. The order of the polynomial was taken to be 20.

The scaled LDFs for the numerically calculated frequencies of system-size earthquakes are shown in Fig. 5. The system size  $L$  varies as  $L = 6, \dots, 12, 16$ , and  $32$ . The LDFs in the cases of  $L = 16$  and  $32$  are obtained by the cloning Monte Carlo method, while the others are by the matrix method. It is difficult to calculate the LDF for MA with  $f = 1.0$  and  $L = 16$  because the number of system-size earthquakes is very small, and so instead, the LDF for  $L = 13$  is plotted. The red solid curve in each panel denotes the LDF of the Poisson process  $\phi_P(z)$  (20), the blue dashed curve denotes the scaled LDF of the one-site system  $\phi_1(z)$ , and the black dashed curve denotes the quadratic function as a reference. In calculating  $\phi_1(x)$ , we introduced the effective triggering rate  $f'$  of the system-size earthquake, which is assumed to be proportional to the number of trigger sites  $n_t$ , and so  $f' = n_t f$ . When the LDF is a quadratic function,  $P(x)$  is Gaussian. It is noted that  $P(x)$  becomes Gaussian when the earthquake sequence is i.i.d. Scaling by  $x_m$  worked well for some cases. The scaled LDFs for different values of  $L$  collapse onto a curve for M1 with  $f = 0.01$  (Fig. 5(c)) and for MA with  $f = 1.0$  (Fig. 5(j)). In contrast, in the other cases, the LDFs for different values of  $L$  are scattered.

From the comparison of the obtained LDFs of the system-size earthquakes with those of the Poisson process, we find three phases of behavior: Poisson, half-Poisson, and non-Poisson. The LDFs for MA with  $f = 1.0$  are well approximated by the LDF of the Poisson process, independent of  $L$  (Fig. 5(j)); this is called the ‘Poisson phase’. In this case, the occupation of all sites rarely occurs because of frequent triggering. For M1, M2, and M3, with  $f = 0.01$ ,  $\phi_P(z)$  fairly approximates the simulated LDFs for  $z < 0$ , while it deviates from the LDFs for  $z > 0$  (Figs. 5(c), (f), and (i)). We call this behavior the ‘half-Poisson phase’. In the half-Poisson phase, the plots are better approximated by  $\phi_1(z)$  than by  $\phi_P(z)$ .  $\phi_1(z)$  is close to  $\phi_P(z)$  for  $z < 0$ , though it deviates from  $\phi_P(z)$  for  $z > 0$ .  $\phi_1(z)$  gives a good approximation in the half-Poisson phase. This is because earthquakes of  $s < L$  rarely occur for small  $f$ , and the system-size earthquakes are dominant. When the frequency of system-size earthquakes is much larger than that of other earthquakes, the system is approximated by the transition between two states  $\{0, \dots, 0\}$  and  $\{1, \dots, 1\}$ , which is similar to the one-site forest-fire model of two states  $\{0\}$  and  $\{1\}$ . For  $z < 0$ , the number of system-size earthquakes during a time interval  $t$  is small compared to the mean, and the time intervals between successive system-size earthquakes are longer than the average. When the time interval is long, all the sites tend to be loaded before triggering. This results in nearly random occurrence of system-size earthquakes, and the LDF may be approximated by the LDF of the Poisson process. It is remarked that the origin of the Poisson behavior in the half-Poisson phase for  $z < 0$  is different from that of the Poisson phase. The Poisson behavior in the half-Poisson phase originates in the Poisson process of triggering, while in the Poisson phase it originates in the random occurrence of full loading of the system.

For the other cases, the simulated LDFs clearly deviate from that of the Poisson process, and this is called the ‘non-Poisson phase’. Although the LDFs show non-Poisson behavior for MA with  $f = 0.1$  and  $0.01$  (Figs. 5(k) and (l)), we anticipate that the distributions converge to the Poisson curve with increasing  $L$ , as discussed in the next section. For this reason, we write N(P) for MA with  $f = 0.1$  and  $0.01$ . The phases of the LDFs thus classified are shown in Table II. For M1, M2, and M3, we do not have a clear explanation for the non-Poisson behavior. The troughs in the LDFs for M1 with  $f = 1.0$  and  $0.1$  (Figs. 5(a) and (b)), and for M2 and M3 with  $f = 0.1$  (Figs. 5(e) and (h)), are deeper for larger  $L$ . In contrast, the troughs in the LDFs for M2 and M3 with  $f = 1.0$  (Figs. 5(d) and (g)) are shallower for larger  $L$ . A deep trough in the LDF indicates that the frequency of system-size earthquakes is commonly close to  $x_m$ , and as the trough becomes deeper, the system-size earthquakes occur more periodically because the earthquake sequence is a renewal process. A renewal process is characterized by a distribution of time intervals between successive events (here the events are the system-size earthquakes). The sequence of the system-size earthquakes in the present forest-fire models is a renewal process because the system always becomes empty after each system-size earthquake. The LDF for the frequency of a renewal process depends on two factors: the time-interval distribution and the trajectory of the time intervals (see the appendix). Let us assume an LDF which has a very deep trough at  $x_m$ , and a distribution of time intervals between successive system-size earthquakes with the dominant peak at  $\Delta t$ . In this case,  $P(x) (\sim e^{-t\phi(x)})$  is close to a delta function with the peak at  $x_m$ , which means that the probability of the trajectories of system-size earthquakes deviating from  $x = x_m$  is very small. The trajectories that give  $x = x_m$



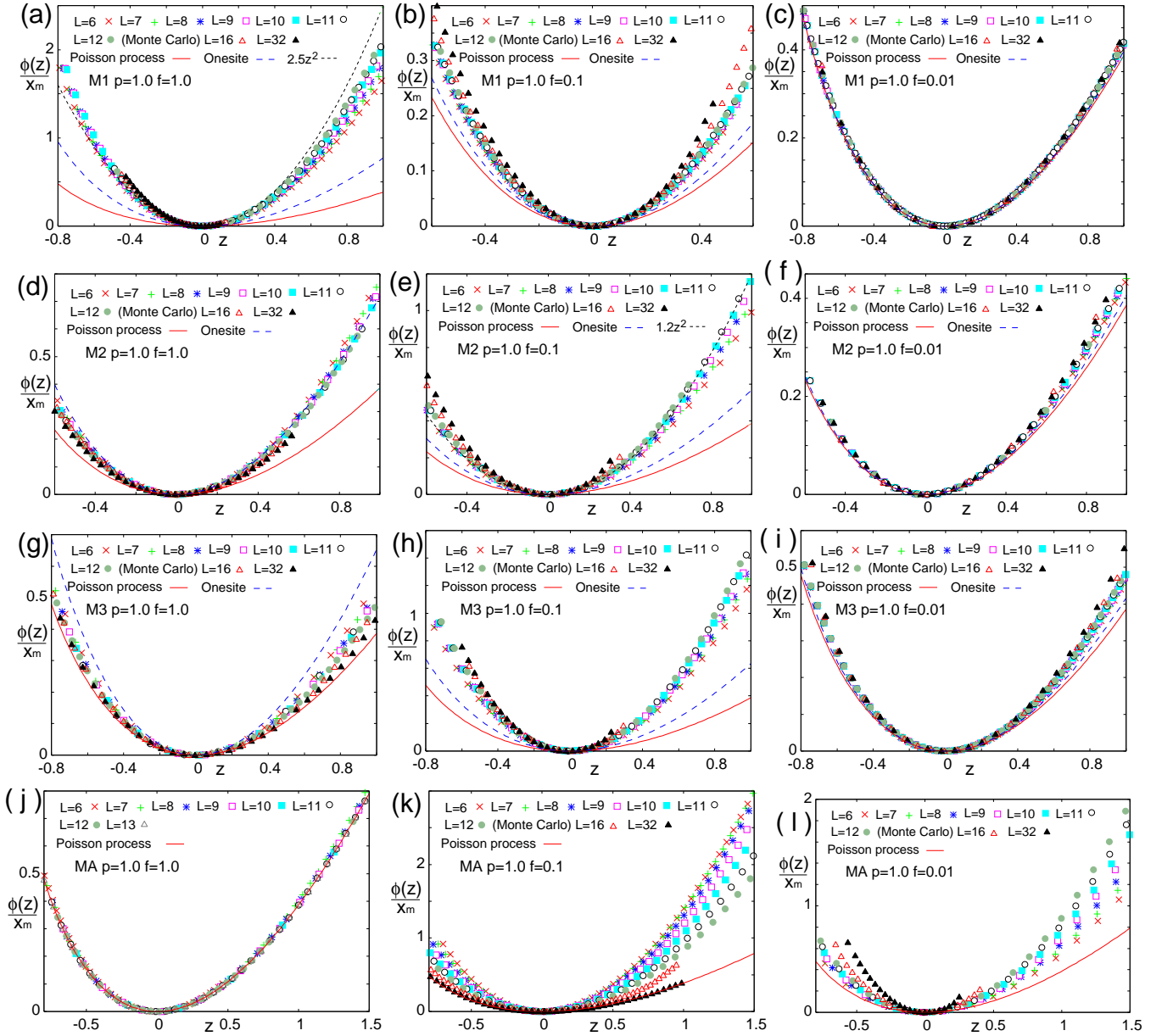


FIG. 5: The plot of  $\phi(z)$  for models M1 ( $f = 1.0$ (a),  $0.1$ (b),  $0.01$ (c)), M2 ( $f = 1.0$ (d),  $0.1$ (e),  $0.01$ (f)), M3 ( $f = 1.0$ (g),  $0.1$ (h),  $0.01$ (i)) and MA ( $f = 1.0$ (j),  $0.1$ (k),  $0.01$ (l)).

consist of almost equal time intervals  $\sim \Delta t$  or of time intervals that contain deviations from  $\Delta t$ . The latter case is unlikely to occur because a deviation from  $\Delta t$  must be supplemented by other earthquakes that also deviate from  $\Delta t$ .

## IV. DISCUSSION AND CONCLUSIONS

### A. The size-frequency distribution and the LDF of system-size earthquakes

The LDFs for MA with  $f = 0.1$  (Fig. 5(k)) and  $0.01$  (Fig. 5(l)), which belong to N(P), display a clear failure of the scaling by the mean frequency  $x_m$ . For MA with  $f = 1.0$ , (i) the LDF shows the Poisson phase (Fig. 5(j)); (ii) there is no peak at the system-size earthquake in the size-frequency distribution (Fig. 3 (d)); and (iii) the mean frequency

	$f=1.0$	$f=0.1$	$f=0.01$
M1	N	N	H
M2	N	N	H
M3	N	N	H
MA	P	N(P)	N(P)

TABLE II: Classification of the LDF by  $f$  of the four models. P, N, and H denote the Poisson, non-Poisson, and the half-Poisson. The N(P) phase becomes Poisson when  $L$  is large enough.

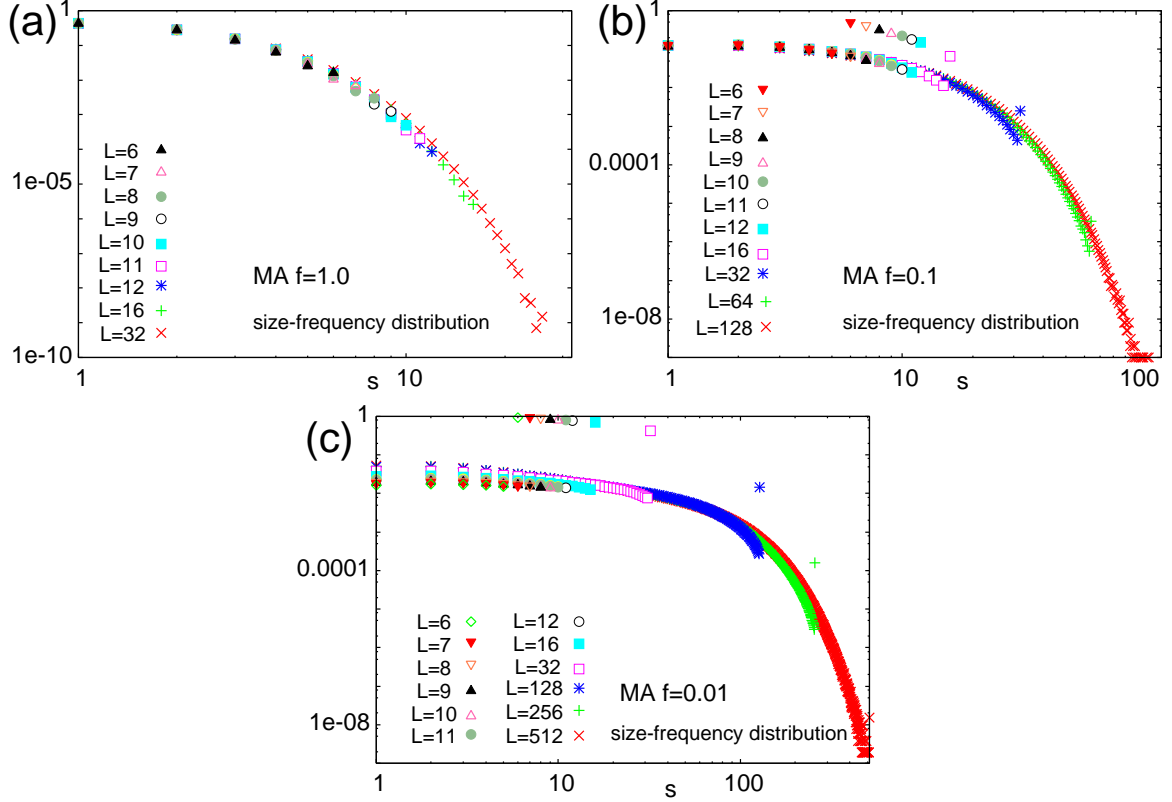


FIG. 6: Size-frequency distributions of MA for different  $L$  with (a)  $f = 1.0$ , (b)  $f = 0.1$ , and (c)  $f = 0.01$ .

of system-size earthquakes is approximated by an exponential function of  $L$  (Fig. 4(b)). The characteristics (ii) and (iii) hold for  $f = 0.1$  and  $0.01$  when the system size is large enough, as we show in detail below.

Fig. 6 shows the size-frequency distributions of earthquakes for MA with  $f = 1.0, 0.1$ , and  $0.01$ . The distributions for MA with  $f = 1.0$  do not show peaks at  $s = L$ , while those for  $f = 0.1$  do show peaks at  $s = L$ . For MA with  $f = 0.1$  and  $0.01$ , the peak of frequency at  $s = L$  is clear when  $L$  is small, but it becomes unclear with an increase in  $L$ . This suggests that the characteristic (ii) is satisfied for large  $L$ .

Fig. 4(b) shows that the mean frequency of system-size earthquakes behaves asymptotically as an exponential function of  $L$  for MA with large  $L$ , indicating that the characteristic (iii) holds for large  $L$ . Thus we confirm that the characteristics (ii) and (iii) are satisfied for large  $L$  and MA with  $f = 0.1$  and  $0.01$ , and we expect that the characteristic (i) is also satisfied in these cases.

In the model MA with large  $L$  and  $f$ , earthquakes are frequently triggered before the system is fully loaded. This results in the occurrence of many small earthquakes and few system-size earthquakes, leading to characteristics (ii) and (iii). The configuration of fully loaded system appears rarely and randomly, and therefore the sequence of the system-size earthquake follows approximately the Poisson process. Thus we expect that the LDFs of system-size earthquakes in the model MA may be categorized into the Poisson phase for large  $L$ . Note that this expectation has not been proven theoretically; it is left for a future study.

So far we have discussed relations among the LDF for the frequency of system-size earthquakes, the size-frequency distribution of earthquakes, and the dependence on system size of the mean frequency of system-size earthquakes for the model MA, which suggest that the characteristic time sequence of system-size earthquakes is related to the size-frequency distribution. What about the other models? For the non-Poisson behaviors of M1, M2, and M3, we do not find clear relations between the size-frequency distributions and the LDFs for the frequency of system-size earthquakes. For M3 with  $f = 1.0$ , with increasing  $L$ , the LDF approaches the LDF of the Poisson process (Fig. 5(g)), which may be similar to the N(P) phase. However, in this case, neither the characteristic (ii) nor (iii) is satisfied, suggesting that the behavior is different from that of the model MA.

## B. Conclusion

The size-frequency distributions and the LDFs for the frequency of system-size earthquakes for four kinds of 1D forest-fire models were calculated numerically. The LDFs for the frequency of the system-size earthquakes mostly deviate from the LDF of the Poisson process. We classified the behaviors of the size-frequency distributions into three phases: supercritical, critical, and subcritical. We also classified the behaviors of the LDFs for the frequency of the system-size earthquakes into three phases: Poisson, non-Poisson, and half-Poisson. The Poisson phase of the LDF is related to the subcritical phase, where the peak at the system-size earthquake is unclear in the size-frequency distribution. This relation has yet to be confirmed by the calculation of the LDF for large  $L$  where the mean frequency of system-size earthquakes exponentially decreases with an increase in system size. The calculation of the LDF for such a large  $L$  is difficult at present.

For real seismic activities, the statistical properties of the frequency of large earthquakes of magnitudes greater than 7.0 was studied in [34], and it was found that the sequence of large earthquakes obeys the Poisson process for magnitudes over 7.3. Although the 1D forest-fire models are too simple to characterize the complex properties of real earthquakes, the present study may provide some insight into the relationship between the size-frequency distribution of earthquakes and the statistical properties of interevent times of system-size earthquakes. Whether the system-size earthquakes occur at random, obeying the Poisson process, or with some regularity, is related to fault properties.

The present study focused on earthquakes. Deviation from the Poisson process appears in many contexts of science, such as bunching in the photon-counting process and a spike train of neurons. The LDF approach can also be applied to those phenomena, and we hope this study furthers understanding of the underlying physics of systems described by point processes.

## Acknowledgments

This research is supported by the MEXT project “Evaluation and disaster prevention research for the coming Tokai, Tonankai and Nankai earthquakes”. The numerical calculations in this study was partly carried out at YITP in Kyoto University.

## Appendix A: LDF for the frequency of a renewal process

The sequence of system-size earthquakes in the 1D forest-fire models can be regarded as a renewal process because each system-size earthquake releases all the stress from the system, leading to an empty configuration. Note that the statistics of the occupation time of a renewal process is discussed in [35].

We consider the probability  $p(N, t)$  that  $N$  events occur during a time interval  $t$ .  $p(N, t)$  can be regarded as a function with frequency  $x = N/t$ , and the function can be written as  $P(x)$ . The LDF for the frequency  $\phi(x)$  behaves asymptotically as the probability  $P(x)$ , which is expressed by

$$\phi(x) = \lim_{t \rightarrow \infty} -\frac{1}{t} \log P(x). \quad (\text{A1})$$

We consider the sequence of interval times between successive events  $t_1, \dots, t_{N-1}$ , and the probability density function  $w(t_i)$ . For a renewal process, the probability of earthquakes occurring with intervals  $t_1, \dots, t_{N-1}$  is written as

$$\text{Prob}(t_1, \dots, t_{N-1}) = \prod_{i=1}^{N-1} w(t_i). \quad (\text{A2})$$

Since  $Prob(t_1, \dots, t_{N-1})$  is the probability for the trajectory of a sequence of earthquakes, we need to sum up the probabilities of all the trajectories to obtain the probability of  $N$  earthquakes in the time interval  $t$ . Thus  $p(N, t)$  is given by the integral with  $t_1, \dots, t_N$  of  $Prob(t_1, \dots, t_N)$  times the probability of no earthquakes during  $t - \sum_{i=1}^{N-1} t_i$  for  $t \geq \sum_{i=1}^{N-1} t_i$ .  $p(N, t)$  is written as

$$p(N, t) = \int_0^t dt_{N-1} \cdots \int_0^t dt_1 \prod_{i=1}^{N-1} [w(t_i)] \psi(t - \sum_{i=1}^{N-1} t_i) \Theta(t - \sum_{i=1}^{N-1} t_i), \quad (\text{A3})$$

where  $\Theta(t)$  is a step function satisfying  $\Theta(t) = 1$  for  $t \geq 0$  and  $\Theta(t) = 0$  otherwise.  $\Theta(t - \sum_{i=1}^{N-1} t_i)$  selects the trajectories between  $t \geq \sum_{i=1}^{N-1} t_i$  and  $t < \sum_{i=1}^{N-1} t_i$ .  $\psi(t)$  is the probability of no earthquakes during  $t$ . Here we note that

$$\psi(t) = \int_t^\infty dt' w(t') \quad (\text{A4})$$

is satisfied.

From the equations (A3) and (A4), we see that  $p(N, t)$  is determined by the two factors:  $w(t)$  and the trajectories of the time intervals  $t_1, \dots, t_{N-1}$  that satisfy  $t \geq \sum_{i=1}^{N-1} t_i$ .

- 
- [1] B. Gutenberg and C. F. Richter, Bull. Seismol. Soc. Am. **34**, 185 (1994).
  - [2] I. Main, Rev. Geophys. **34**, 433 (1996).
  - [3] H. Touchette, Phys. Rep. **478**, 1 (2009).
  - [4] D. Sornette, *Critical Phenomena in Natural Sciences* 2nd. ed. p. 59 (Springer-Verlag, Berlin, 2006).
  - [5] B. Derrida, J. Stat. Mech. P07023 (2007).
  - [6] T. Nemoto and S. I. Sasa, Phys. Rev. E **83**, 030105(R) (2011); **84**, 061113 (2011).
  - [7] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
  - [8] J. P. Garrahan and I. Lesanovsky, Phys. Rev. Lett. **104**, 160601 (2010).
  - [9] A. A. Budini, Phys. Rev. E **84**, 011141 (2011).
  - [10] J. Li, Y. Liu, J. Ping, S. S. Li, X. Q. Li and Y. J. Yan, Phys. Rev. B **84**, 115319 (2011).
  - [11] Y. Ogata, Jour. Am. Stat. As. **83**, 9 (1988).
  - [12] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
  - [13] V. Lecomte and J. Tailleur, J. Stat. Mech. P03004 (2007).
  - [14] C. Giardinà, J. Kurchan, V. Lecomte and J. Tailleur, J. Stat. Phys. **145**, 787 (2011).
  - [15] P. Bak, K. Chen, and C. Tang, Phys. Lett. A **147**, 297 (1990).
  - [16] B. Drossel and F. Schwabl, Phys. Rev. Lett. **69**, 1629 (1992).
  - [17] C. L. Henley, Phys. Rev. Lett. **71**, 2741 (1993).
  - [18] M. Paczuski and P. Bak, Phys. Rev. E **48**, R3214 (1993).
  - [19] M. R. Yoder, D. L. Turcotte and J. B. Rundle, Phys. Rev. E **83**, 046118 (2011).
  - [20] D. L. Turcotte, Rep. Prog. Phys. **62**, 1377 (1999); Phys. Earth. Planet. Inter. **111**, 275 (1999).
  - [21] M. Otsuka, ZISIN. Ser.II (in Japanese) **24**, 215 (1971).
  - [22] W. I. Newman and D. L. Turcotte, Nonlin. Proc. Geophys. **9**, 453 (2002).
  - [23] M. Vázquez-Prada, Á. González, J. B. Gómez and A. F. Pacheco, Nonlin. Proc. Geophys. **9**, 513 (2002).
  - [24] A. Tejedor, J. B. Gómez and A. F. Pacheco, Phys. Rev. E **79**, 046102 (2009).
  - [25] B. Drossel, S. Clar and F. Schwabl, Phys. Rev. Lett. **71**, 3739 (1993).
  - [26] A. Schadschneider, D. Chowdhury, and K. Nishinari, *Stochastic Transport in Complex Systems*, p. 33 (Elsevier, Amsterdam, 2011).
  - [27] B. Derrida, M. R. Evans, V. Hakeem, and V. Pasquier, J. Phys. A: Math. Gen. **26**, 1493 (1993).
  - [28] T. Sasamoto, J. Phys. A: Math. Gen. **32**, 7109 (1999).
  - [29] N. Rajewsky, L. Santen, A. Schadschneider and M. Schreckenberg, J. Stat. Phys. **92**, 151 (1998).
  - [30] A. Honecker and I. Peschel, Physica A **229**, 478 (1996).
  - [31] T. Mitsudo and S. Takesue, J. Phys. Soc. Jap. **80**, 114001 (2011).
  - [32] T. Mori and H. Kawamura, J. Geophys. Res. **113**, B06301 (2008).
  - [33] H. Kawamura, T. Hatano, N. Kato, S. Biswas and B. K. Chakrabarti, Rev. Mod. Phys. **84**, 839 (2012).
  - [34] E. G. Daub, E. Ben-Naim, R. A. Guyer and P. A. Johnson, Geophys. Res. Lett. **39**, L06308 (2012).
  - [35] C. Godrèche and J. M. Luck, J. Stat. Phys. **104**, 489 (2001).